


```

FT DISULFID 385 418 BY SIMILARITY.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 156 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 856 AA; 97212 MW; 6FAB16AF85107FE0 CRC64;

```

Query Match 83.0%; Score 112; DB 1; Length 856;
 Best Local Similarity 85.7%; Pred. No. 1.5e-08;
 Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

OY 1 RVIRVGRACRAIRHVRIRRGRLRL 28
    ||| ||| ||| ||| ||| ||| ||| |||
Db 828 RVIEVVGACRAIRHVRIRRGRLRL 855

```

```

RESULT 2
ENV_HV1LM STANDARD; PRT; 856 AA.
AC 070626;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide gp160 precursor [contains: Exterior membrane
  glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (LM12.3 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
ON NCBI_TaxID=82834;
RX MEDLINE=95127297; PubMed=7826699;
RA Reitz M.S. Jr., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.M.,
  Shaw G.M., Kong L.I., Weiss S.H., Waters D., Gallo R.C., Blattner W.;
  "Viral variability and serum antibody response in a laboratory worker
  infected with HIV type 1 (HIV type IIB).";
RT AIDS Res. Hum. Retroviruses 10:1143-1155(1994).

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

```

CC -----
DR EMBL: U12055; AAA76690.1; -
DR Glycosub: 070626; -
DR InterPro: IPR000328; Env GP41.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;
  signal.
FT SIGNAL 1 30
FT CHAIN 31 511
FT 512 856
FT DISULFID 54 74
FT DISULFID 119 205
FT DISULFID 126 196
FT DISULFID 131 157
FT DISULFID 218 247
FT DISULFID 228 239
FT DISULFID 296 331
FT DISULFID 378 445
FT DISULFID 385 418
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 156 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 856 AA; 96938 MW; 0C241332CFE6687 CRC64;

```

Query Match 83.0%; Score 112; DB 1; Length 856;
 Best Local Similarity 85.7%; Pred. No. 1.5e-08;
 Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

OY 1 RVIRVGRACRAIRHVRIRRGRLRL 28
    ||| ||| ||| ||| ||| ||| ||| |||
Db 828 RVIEVVGACRAIRHVRIRRGRLRL 855

```

```

RESULT 3
ENV_HV1BR STANDARD; PRT; 861 AA.
AC P03377;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide gp160 precursor [contains: Exterior membrane
  glycoprotein (GP120); Transmembrane glycoprotein (GP41)].

```

[illegible]

50	SEQUENCE	861 AA;	97487 MM;	04DE2B4D4E4PD63A CRC64;
	Query Match		83.0%;	Score 112; DB 1; Length 861;
	Best Local Similarity	85.7%;	Pred. No. 1.5e+08;	
	Matches 24;	Conservative 0;	Mismatches 4;	Indels 0; Gaps 0;
Qy	1	RVIRVORACRAIRHIVRIRROGLRIL	28	
Db	833	RVIEVVGACRAIRHPRIRROGLERIL	860	
	RESULT	4		
	ID	ENV_HV1H3	STANDARD;	PRT; 856 AA.
AC	P04624;			
DT	13-AUG-1987 (Rel. 05, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	15-Jun-1999 (Rel. 38, Last annotation update)			
DE	Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].			
DE	ENV.			
OS	Human immunodeficiency virus type 1 (HXB3 isolate) (HIV-1).			
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.			
OX	NCBI_TaxID=11707;			
RA	SEQUENCE FROM N.A.			
RA	MEDLINE=8528248; PubMed=298795;			
RA	Crowl R., Ganguly R., Gordon M., Conroy R., Schaber M., Kramer R., Shaw G.M., Wong-Staal F., Reddy E.P.;			
RT	"HIV-III env gene products synthesized in E. coli are recognized by antibodies present in the sera of AIDS patients."			
RL	Cell 41:979-986(1985).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; M14100; AAA44679.1; -.			
DR	HIV; M14100; ENVSHXB3.			
DR	InterPro: IPR000328; Env_GP41.			
DR	InterPro: IPR000777; GP120.			
DR	Pfam: PF00516; GP120: 1.			
DR	Pfam: PF00517; GP41: 1.			
KW	AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal.			
FT	SIGNAL	1	30	
FT	CHAIN	512	856	EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN	512	856	TRANSMEMBRANE GLYCOPROTEIN.
FT	DISULFID	119	205	BY SIMILARITY.
FT	DISULFID	126	196	BY SIMILARITY.
FT	DISULFID	131	157	BY SIMILARITY.
FT	DISULFID	218	247	BY SIMILARITY.
FT	DISULFID	228	239	BY SIMILARITY.
FT	DISULFID	296	331	BY SIMILARITY.
FT	DISULFID	378	445	BY SIMILARITY.
FT	DISULFID	385	418	BY SIMILARITY.
FT	CARBOHYD	88		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	136	136	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	141	141	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	156	156	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	160	160	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	186	186	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	197	197	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	230	230	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	234	234	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	241	241	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	262	262	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	276	276	N-LINKED (GLCNAC. . .) (POTENTIAL).

```
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 856 AA; 97188 MW; 3373C6BB84C1AFC CRC64;
```

Query Match Score 76.3%; DB 1; Length 856;

Best Local Similarity 82.1%; Pred. No. 2.9e-07; Mismatches 5; Indels 0; Gaps 0;

Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 RVIRVORACRAIRHIVRIROGLRRL 28

Db 828 RVIEVQEAIRHIVRIROGLRRL 855

```
RESULT 5
ENV_HV1ND STANDARD; PRT; 846 AA.
AC P18799;
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (NDK isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11695;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90034200; PubMed=2806917;
RA Spire B., Sire J., Zacher V., Rey F., Barre-Sinoussi F., Galibert F.,
RA Hampe A., Chermann J.C.;
RT "Nucleotide sequence of HIV-1 NDK: a highly cytopathic strain of the
RT human immunodeficiency virus.";
RL Gene 81:275-284(1989).
CC -1- MISCELLANEOUS: NDK, ISOLATED FROM A ZAIRIAN PATIENT AFFECTED WITH
CC AIDS. AND IS A HIGHLY CYTOPATHOGENIC STRAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M27323; AAA44873.1; -.
CC PIR: J00066; VCLJND.
CC HIV: M27323; ENVSNDK.
CC InterPro: IPR000328; Env.GP41.
CC InterPro: IPR000777; GP120.
CC Pfam: PF00517; GP120; 1.
CC Pfam: PF00517; GP41; 1.
CC AIDS: Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW Signal.
FT SIGNAL 1 29
```

```
FT CHAIN 30 501 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 502 846 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 200 BY SIMILARITY.
FT DISULFID 125 191 BY SIMILARITY.
FT DISULFID 130 152 BY SIMILARITY.
FT DISULFID 213 242 BY SIMILARITY.
FT DISULFID 223 234 BY SIMILARITY.
FT DISULFID 291 328 BY SIMILARITY.
FT DISULFID 374 435 BY SIMILARITY.
FT DISULFID 381 408 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 182 182 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 229 229 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 236 236 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 257 257 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 284 284 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 351 351 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 388 388 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 395 395 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 438 438 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 451 451 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 452 452 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 601 601 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 606 606 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 615 615 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 627 627 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 846 AA; 96476 MW; 8A3BD5A27DE2EB3 CRC64;
```

Query Match Score 75.6%; DB 1; Length 846;

Best Local Similarity 75.0%; Pred. No. 4e-07; Mismatches 4; Indels 0; Gaps 0;

Matches 21; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 RVIRVORACRAIRHIVRIROGLRRL 28

Db 818 RVIEVQEAIRHIVRIROGLRRL 845

```
RESULT 6
ENV_HV1W2 STANDARD; PRT; 847 AA.
AC P05880;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (WMJ2 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11705;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86235450; PubMed=3012778;
RA Hahn B.H., Shaw G.M., Taylor M.E., Redfield R.R., Markham P.D.,
RA Salahuddin S.Z., Wong-Staal F., Gallo R.C., Parks E.S., Parks W.P.;
RT "Genetic variation in HTLV-III/LAV over time in patients with AIDS or
RT at risk for AIDS.";
RL Science 232:1348-1353(1986).
CC -1- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM
CC BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO
CC WAS PERINATALLY INFECTED BY HER MOTHER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```


FT	CARBOHYD	814	814	N-LINKED (GLCNAC. .) (POTENTIAL).
SO	SEQUENCE	853 AA;	96912 MW;	3377B993B6F22ABA CMC64;
	Query Match		74.8%;	Score 101; DB 1; Length 853;
	Best Local Similarity	82.1%;	Pred. No. 5.6e-07;	
	Matches 23;	Conservative 0;	Mismatches 5;	Indels 0; Gaps 0;
QY	1 RVIRVQAPCAIRHIVRRIRROGLRRIL 28			
Db	826 RVIEVQAGAYRAIRHPRIRROGLRRIL 853			
RESULT	8			
ENV_HV1B1	STANDARD;	PRT;	856 AA.	
AC	P03375;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Envelope polypeptide gp160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].			
GN	ENV.			
OS	Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1).			
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.			
OX	NCBI_Taxid=11678;			
RN	[1]			
RX	SEQUENCE FROM N.A.			
RX	MEDLINE=8511123; PubMed=2578615;			
RA	Ratner L., Hesselhine W., Patarca R., Liyak K.J., Starcich B.R.,			
RA	Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,			
RA	Baumeister K., Ivanoff L., Peteway S.R. Jr., Pearson M.L.,			
RA	Latebberger J.A., Papas T.S., Grayeb J., Chang N.T., Gallo R.C.,			
RA	Wong-Staal F.;			
RT	"Complete nucleotide sequence of the AIDS virus, HTLV-III.;"			
RL	Nature 313:277-284(1985).			
RN	[2]			
RN	DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.			
RX	MEDLINE=90285159; PubMed=235506;			
RA	Leonard C.K., Spellman M.W., Riddle L., Harris R.J., Thomas J.N.,			
RA	Gregory T.J.;			
RT	"Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells."			
RL	J. Biol. Chem. 265:10373-10382(1990).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
CC	-----			
DR	EMBL; M15654; AAA44205.1; .			
DR	PIR; A03973; VCLJH3.			
DR	HIV; M15654; ENVSBH102.			
DR	InterPro: IPR000378; Env.GP41.			
DR	InterPro: IPR000777; GP120.			
DR	PIfam; PF00516; GP120; 1.			
DR	PIfam; PF00517; GP41; 1.			
KW	AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;			
KW	Signal.			
FT	SIGNAL	1	30	
FT	CHAIN	31	511	
FT	CHAIN	512	856	EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	DISULFID	54	74	TRANSMEMBRANE GLYCOPROTEIN.
FT	DISULFID	119	205	
FT	DISULFID	126	196	
FT	DISULFID	131	157	
FT	DISULFID	218	247	
FT	DISULFID	228	239	

Query Match	Best Local Similarity	74.8%, Score 101, DB 1, Length 856;
Matches 23; Conservative	0; Mismatches 5; Indels 0; Gaps 0;	
Qy 1 RVIRVORACRAIRHIVRIRGRLRL 28		
Db 828 RVIEVQAGYRAIRHIVPRIRGRLRL 855		
RESULT 9		
ENV_HV1PV STANDARD: PRT: 856 AA.		
ID ENV_HV1PV		
AC P03376;		
DT 21-JUL-1986 (Rel. 01, Created)		
DT 21-JUL-1986 (Rel. 01, Last sequence update)		
DT 15-JUL-1999 (Rel. 38, Last annotation update)		
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].		
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41).		
GN ENV.		
OC Human immunodeficiency virus type 1 (PV22 isolate) (HIV-1).		
OC Viruses; Retroid viruses; Retroviridae; Lentiviruses.		
OC NCBI_TaxID=11700;		
OX		
RN [1]		
RP SEQUENCE FROM N.A.		
RA MEDLINE=8511157; PubMed=2982104;		
RA Muesing M.A., Smith D.H., Cabradilla C.D., Benton C.V., Lasky L.A.,		
RA Capon D.J.;		
RT "Nucleic acid structure and expression of the human		
RT AIDS/lymphadenopathy retrovirus.";		
RL Nature 313:450-458(1985).		
CC -----		
CC This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC between the Swiss Institute of Bioinformatics and the EMBL outstation		
CC the European Bioinformatics Institute. There are no restrictions on its		
CC use by non-profit institutions as long as its content is in no way		
CC modified and this statement is not removed. Usage by and for commercial		

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: K02083; AAB59873.1; ALT_SEQ.

DR EMBL: X01762; CAA25903.1; ALT_SEQ.

DR PIR: A03974; VCLJ3V.

DR HIV: K02083; ENVSPV22.

DR InterPro: IPR000328; Env_GP41.

DR InterPro: IPR000777; GP120.

DR Pfam: PF00516; GP120; 1.

DR Pfam: PF00517; GP41; 1.

KM Aids: Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal.

FT SIGNAL 1 30

FT CHAIN 31 511

FT CHAIN 512 856

FT DISULFID 54 74

FT DISULFID 119 205

FT DISULFID 126 196

FT DISULFID 131 157

FT DISULFID 218 247

FT DISULFID 228 239

FT DISULFID 296 331

FT DISULFID 378 445

FT DISULFID 385 418

FT CARBOHYD 88 88

FT CARBOHYD 136 136

FT CARBOHYD 141 141

FT CARBOHYD 156 156

FT CARBOHYD 160 160

FT CARBOHYD 186 186

FT CARBOHYD 197 197

FT CARBOHYD 230 230

FT CARBOHYD 234 234

FT CARBOHYD 241 241

FT CARBOHYD 262 262

FT CARBOHYD 276 276

FT CARBOHYD 289 289

FT CARBOHYD 295 295

FT CARBOHYD 301 301

FT CARBOHYD 332 332

FT CARBOHYD 339 339

FT CARBOHYD 356 356

FT CARBOHYD 366 366

FT CARBOHYD 392 392

FT CARBOHYD 397 397

FT CARBOHYD 406 406

FT CARBOHYD 448 448

FT CARBOHYD 463 463

FT CARBOHYD 611 611

FT CARBOHYD 616 616

FT CARBOHYD 625 625

FT CARBOHYD 637 637

FT CARBOHYD 674 674

FT CARBOHYD 750 750

FT CARBOHYD 816 816

FT CARBOHYD 856 AA; 97339 MW; 5FCBD1DC3C1209B3 CRC64;

SO SEQUENCE

Query Match 74.8%; Score 101; DB 1; Length 856;
Best Local Similarity 82.1%; Pred. No. 5.6e-07;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 RVIRVQACRAIRHIVRIROGLRRIL 28
DB 828 RVIEVQAGYRAIRHIVRIROGLRRIL 855

RESULF 10
ENV_HV1W1
ID ENV_HV1W1 STANDARD; PRT; 856 AA.
AC P31872;
DT 01-JUL-1993 (rel. 26, created)

DT 01-JUL-1993 (rel. 26, last sequence update)

DT 15-JUL-1999 (rel. 38, last annotation update)

DE Identification and characterization of conserved and variable regions in the envelope gene of HTLV-III/LAV, the retrovirus of Aids."

RT Aids: 45-637-648(1986).

RT Aids: 45-637-648(1986).

RL -1- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO WAS PERINATALLY INFECTED BY HER MOTHER.

CC PIR: A24774; VCLJ3V.

CC InterPro: IPR000328; Env_GP41.

CC InterPro: IPR000777; GP120.

DR Pfam: PF00516; GP120; 1.

DR Pfam: PF00517; GP41; 1.

KM Aids: Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal.

FT SIGNAL 1 29

FT CHAIN 31 510

FT CHAIN 511 856

FT DISULFID 54 73

FT DISULFID 118 205

FT DISULFID 125 196

FT DISULFID 130 152

FT DISULFID 218 247

FT DISULFID 228 239

FT DISULFID 296 330

FT DISULFID 376 444

FT DISULFID 383 417

FT CARBOHYD 87 87

FT CARBOHYD 134 134

FT CARBOHYD 140 140

FT CARBOHYD 151 151

FT CARBOHYD 155 155

FT CARBOHYD 183 183

FT CARBOHYD 197 197

FT CARBOHYD 234 234

FT CARBOHYD 241 241

FT CARBOHYD 262 262

FT CARBOHYD 276 276

FT CARBOHYD 289 289

FT CARBOHYD 295 295

FT CARBOHYD 331 331

FT CARBOHYD 338 338

FT CARBOHYD 354 354

FT CARBOHYD 360 360

FT CARBOHYD 394 394

FT CARBOHYD 394 394

FT CARBOHYD 404 404

FT CARBOHYD 447 447

FT CARBOHYD 459 459

FT CARBOHYD 611 611

FT CARBOHYD 616 616

FT CARBOHYD 625 625

FT CARBOHYD 637 637

FT CARBOHYD 856 AA; 97526 MW; DB68D1E49C404DE9 CRC64;

SO SEQUENCE

Query Match 74.8%; Score 101; DB 1; Length 856;
Best Local Similarity 78.6%; Pred. No. 5.6e-07;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 RVIRVQACRAIRHIVRIROGLRRIL 28

D8	B28	RVIEWORICRAITHIPRRIRGOGLERAL	855
<div style="text-align: center;"> </div>			
<div style="text-align: right;">-----</div>			
RESULT	11		
ENV_HV1B8			
ID ENV_HV1B8	STANDARD;	PRT;	851 AA.
AC P04582:			
DT 13-AUG-1987	(Rel. 05, Created)		
DT 13-AUG-1987	(Rel. 05, Last sequence update)		
DT 15-JUL-1999	(Rel. 38, Last annotation update)		
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].			
DE ENY.			
OS Human immunodeficiency virus type 1 (BH isolate) (HIV-1).			
OC Viruses; Retroid viruses; Retroviridae; Lentivirns.			
OK NCBI_TaxID=11684;			
RN [1]			
RP SEQUENCE FROM N.A.			
RX MEDLINE=8511123; PubMed=2578615;			
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R., Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A., Baumeister K., Ivanoff L., Petenay S.R. Jr., Pearson M.L., Lutenberger J.A., Papas T.S., Grzybowski J., Chang N.T., Gallo R.C., Wong-Staal F.;			
RA "Complete nucleotide sequence of the AIDS virus, HTLV-II.";			
RL Nature 313:277-284(1985).			
CC -----			
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way CC modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC -----			
DR EMBL; K02011; AAAA4661.1; -.			
DR HIV; K02011; ENVSBH8.			
DR GlycoSuiteDB: P04582; -.			
DR InterPro: IPR000328; Env_GP41.			
DR InterPro: IPR000777; GP120.			
DR Pfam: PF00516; GP120; 1.			
DR Pfam: PF00517; GP41; 1.			
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal.			
FT SIGNAL	1	30	
FT CHAIN	31	506	EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN	507	851	TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFD	54	74	BY SIMILARITY.
FT DISULFD	119	205	BY SIMILARITY.
FT DISULFD	126	196	BY SIMILARITY.
FT DISULFD	131	157	BY SIMILARITY.
FT DISULFD	218	247	BY SIMILARITY.
FT DISULFD	228	239	BY SIMILARITY.
FT DISULFD	296	331	BY SIMILARITY.
FT DISULFD	378	440	BY SIMILARITY.
FT DISULFD	385	413	BY SIMILARITY.
FT CARBOHYD	88	88	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT CARBOHYD	136	136	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT CARBOHYD	141	141	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT CARBOHYD	156	156	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT CARBOHYD	160	160	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT CARBOHYD	186	186	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT CARBOHYD	197	197	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT CARBOHYD	230	230	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT CARBOHYD	234	234	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT CARBOHYD	241	241	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT CARBOHYD	262	262	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT CARBOHYD	276	276	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT CARBOHYD	295	295	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT CARBOHYD	301	301	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT CARBOHYD	332	332	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT CARBOHYD	339	339	N-LINKED (GLCNAC . . .) (POTENTIAL).

FT	CARBOHYD	356	356	N-LINKED (GLCNAC. .)	(POTENTIAL).			
FT	CARBOHYD	386	386	N-LINKED (GLCNAC. .)	(POTENTIAL).			
FT	CARBOHYD	392	392	N-LINKED (GLCNAC. .)	(POTENTIAL).			
FT	CARBOHYD	401	401	N-LINKED (GLCNAC. .)	(POTENTIAL).			
FT	CARBOHYD	443	443	N-LINKED (GLCNAC. .)	(POTENTIAL).			
FT	CARBOHYD	458	458	N-LINKED (GLCNAC. .)	(POTENTIAL).			
FT	CARBOHYD	606	606	N-LINKED (GLCNAC. .)	(POTENTIAL).			
FT	CARBOHYD	611	611	N-LINKED (GLCNAC. .)	(POTENTIAL).			
FT	CARBOHYD	620	620	N-LINKED (GLCNAC. .)	(POTENTIAL).			
FT	CARBOHYD	632	632	N-LINKED (GLCNAC. .)	(POTENTIAL).			
FT	CARBOHYD	669	669	N-LINKED (GLCNAC. .)	(POTENTIAL).			
FT	CARBOHYD	745	745	N-LINKED (GLCNAC. .)	(POTENTIAL).			
FT	CARBOHYD	811	811	N-LINKED (GLCNAC. .)	(POTENTIAL).			
SO	SEQUENCE	851 AA;	96644 MM;	D16A3C90857785F1	CRC64;			
Query Match		73.3%;	Score 99;	DB 1;	Length 851;			
Best Local Similarity		78.6%;	Pred. No. 1.1e-06;					
Matches 22;		Conservative 1;	Mismatches 5;	Indels 0;	Gaps 0;			
Oy	1 RVIRVORACRAIRHIVRIRROGLERIL 28							
	: :							
Db	823 RVIETVQAYRAIRIRIPRIRROGLERIL 850							
RESULT 12								
ENV_HV122								
ID ENV_HV122	STANDARD;	PRT;	853 AA.					
AC P12487;								
DT 01-OCT-1989	(Rel. 12, Created)							
DT 01-OCT-1989	(Rel. 12, Last sequence update)							
DT 16-OCT-2001	(Rel. 40, Last annotation update)							
DE Envelope polypeptide	Gp160 precursor [Contigs: Exterior membrane glycoprotein (Gp120); Transmembrane glycoprotein (Gp41)].							
DE glycoprotein (Gp120);	Transmembrane glycoprotein (Gp41)].							
GN ENV.								
OS Human immunodeficiency virus type 1 (Z2/CDC-Z34 isolate) (HIV-1).								
OC Viruses; Retroid	viruses; Retroviridae; Lentiviruses.							
OC NCBI_Taxid=11683;								
OX [1]								
RN SEQUENCE FROM N.A.								
RP Theodore T., Buckler-White A.;								
RA Submitted (NOV-1988)	to the HIV data bank.							
RL								
CC -----								
CC This SWISS-PROT entry is copyright. It is produced through a collaboration								
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -								
CC the European Bioinformatics Institute. There are no restrictions on its								
CC use by non-profit institutions as long as its content is in no way								
CC modified and this statement is not removed. Usage by and for commercial								
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/								
CC or send an email to license@sib-sib.ch).								
CC -----								
DR EMBL; M22639; AAA45370.1; -								
DR HIV; M22639; ENV5Z226.								
DR InterPro; IPR000328; Env_GP41.								
DR InterPro; IPR000777; GP120.								
DR Pfam; PF00516; GP120; 1.								
DR Pfam; PF00517; GP41; 1.								
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;								
Signal.								
FT SIGNAL	1	31	BY SIMILARITY.					
FT CHAIN	32	508	EXTERIOR MEMBRANE GLYCOPROTEIN.					
FT CHAIN	509	853	TRANSMEMBRANE GLYCOPROTEIN.					
FT DISULFID	53	73						


```
FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 185 185 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 235 235 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 242 242 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 338 338 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 353 353 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 462 462 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 608 608 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 613 613 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 622 622 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 634 634 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 853 AA; 97043 MW; 849B08CBAFF7008 CRC64;
```

Query Match Best Local Similarity 72.6%; Score 98; DB 1; Length 853;
Matches 19; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Oy 1 RVRIVQACRAIRHIVRIROGLRRL 28
|||:|||||:|:|||||:|:
Db 825 RVEIYRRACRAVLIPIRIROGLRSL 852

RESULT 13
ENV_HV1EL STANDARD; PRT; 853 AA.
AC P04581;
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.

OS Human immunodeficiency virus type 1 (ELI isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11689;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=86245056; PubMed=2424612;
RX Allison M., Main-Hobson S., Montagnier L., Sonigo P.,
RT "Genetic variability of the AIDS virus: nucleotide sequence analysis
of two isolates from African patients.";
RL Cell 46:63-74(1986).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

CC
DR EMBL: K03454; AAA44329.1; -;
DR EMBL: A07108; CAA00616.1; -;
DR HIV: K03454; ENVSEL1.
DR InterPro: IPR000328; Env_GP41.

DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW Aids; Coat protein; Polypeptide; Glycoprotein; Transmembrane;
KW Signal.
FT CHAIN 1 31
FT SIGNAL 32 508
FT CHAIN 509 853
FT DISULFID 53 73
FT DISULFID 118 206
FT DISULFID 125 197
FT DISULFID 130 154
FT DISULFID 219 248
FT DISULFID 229 240
FT DISULFID 257 330
FT DISULFID 376 442
FT DISULFID 383 416
FT CARBOHYD 87 87
FT CARBOHYD 129 129
FT CARBOHYD 137 137
FT CARBOHYD 143 143
FT CARBOHYD 153 153
FT CARBOHYD 157 157
FT CARBOHYD 183 183
FT CARBOHYD 188 188
FT CARBOHYD 198 198
FT CARBOHYD 225 225
FT CARBOHYD 242 242
FT CARBOHYD 263 263
FT CARBOHYD 277 277
FT CARBOHYD 290 290
FT CARBOHYD 331 331
FT CARBOHYD 353 353
FT CARBOHYD 384 384
FT CARBOHYD 390 390
FT CARBOHYD 394 394
FT CARBOHYD 400 400
FT CARBOHYD 405 405
FT CARBOHYD 406 406
FT CARBOHYD 411 411
FT CARBOHYD 445 445
FT CARBOHYD 458 458
FT CARBOHYD 459 459
FT CARBOHYD 462 462
FT CARBOHYD 608 608
FT CARBOHYD 613 613
FT CARBOHYD 622 622
FT CARBOHYD 634 634
SQ SEQUENCE 853 AA; 96721 MW; F9CD864DAA0D07A5 CRC64;

Query Match Best Local Similarity 71.1%; Score 96; DB 1; Length 853;
Matches 19; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Oy 1 RVRIVQACRAIRHIVRIROGLRRL 28
|||:|||||:|:|||||:|:
Db 825 RVEIYRRACRAVLIPIRIROGLRSL 852

RESULT 14
ENV_HV1A2 STANDARD; PRT; 855 AA.
AC P03378;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (ARV2/SF2 isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11685;

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85090453; PubMed=2578227;
RA Sanchez-Pescador R., Power M.D., Barr P.J., Steimer K.S.,
RA Stempien M.M., Brown-Shimer S.L., Gee W.W., Renard A., Randolph A.,
RA Levy J.A., Dina D., Luciw P.A.;
RT "Nucleotide sequence and expression of an AIDS-associated retrovirus
RT (ARV-2).";
RL Science 227:484-492(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K02007; AAB59882.1; -.
DR PIR; A03976; VCLJ2.
DR HIV; K02007; ENV5SF2.
DR InterPro: IPR000328; Env_GPA1.
DR InterPro: IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GPA1; 1.
DR AIDs; Coat protein; Polypotein; Glycoprotein; Transmembrane;
KM Signal.
FT SIGNAL 1 29
FT CHAIN 30 509
FT FT 510 855
FT DISULFID 53 73
FT DISULFID 118 208
FT DISULFID 125 199
FT DISULFID 130 155
FT DISULFID 221 250
FT DISULFID 231 242
FT DISULFID 299 333
FT DISULFID 380 442
FT DISULFID 387 415
FT CARBOHYD 87 87
FT CARBOHYD 129 129
FT CARBOHYD 140 140
FT CARBOHYD 154 134
FT CARBOHYD 158 158
FT CARBOHYD 184 184
FT CARBOHYD 190 190
FT CARBOHYD 200 200
FT CARBOHYD 223 233
FT CARBOHYD 244 244
FT CARBOHYD 265 265
FT CARBOHYD 279 279
FT CARBOHYD 292 292
FT CARBOHYD 298 298
FT CARBOHYD 304 304
FT CARBOHYD 334 334
FT CARBOHYD 341 341
FT CARBOHYD 358 358
FT CARBOHYD 364 364
FT CARBOHYD 388 388
FT CARBOHYD 394 394
FT CARBOHYD 400 400
FT CARBOHYD 408 408
FT CARBOHYD 445 445
FT CARBOHYD 458 458
FT CARBOHYD 461 461
FT CARBOHYD 610 610
FT CARBOHYD 615 615
FT CARBOHYD 624 624
FT CARBOHYD 636 636
FT CARBOHYD 815 815
FT CARBOHYD 815 815
SO SEQUENCE 855 AA; 97438 MW; A3BC20573AAC41A2 CRC64;

```

```

Query Match 69.6%; Score 94; DB 1; Length 855;
Best Local Similarity 75.0%; Pred. No. 5.6e-06;
Matches 21; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
QY 1 RTIRVQACRAIRHVRIRIGLRRL 28
DB 827 RTIEVAQRAYRAIRHRRIRIGLERLL 854
RESULT 15
EN ENV_HV1BN STANDARD; PRT: 852 AA.
AC P12488;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (1BR isolate) (HIV-1).
OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11693;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89085613; PubMed=2789516;
RA Anand R., Thayer R., Srinivasan A., Nayyar S., Gardner M., Luciw P.,
RA Dandekar S.;
RT "Biological and molecular characterization of human immunodeficiency
RT virus (HIV-1BR) from the brain of a patient with progressive
RT dementia.";
RL Virology 168:79-89(1989).
CC -1- MISCELLANEOUS: THIS VIRUS IS CYTOPATHICALLY ACTIVE AND WAS
CC HARVESTED FROM THE BRAIN TISSUE OF A NEUROLOGICAL AIDS PATIENT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M21098; AAA44221.1; -.
DR PIR; A31667; VCLJBR.
DR HIV; M21098; ENV5BRVA.
DR InterPro: IPR000328; Env_GPA1.
DR InterPro: IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GPA1; 1.
DR AIDs; Coat protein; Polypotein; Glycoprotein; Transmembrane;
KM Signal.
FT SIGNAL 1 30
FT CHAIN 31 507
FT FT 508 852
FT DISULFID 54 74
FT DISULFID 119 205
FT DISULFID 126 196
FT DISULFID 131 155
FT DISULFID 218 247
FT DISULFID 228 239
FT DISULFID 296 330
FT DISULFID 376 439
FT DISULFID 383 412
FT DISULFID 49 49
FT CARBOHYD 88 88
FT CARBOHYD 135 135
FT CARBOHYD 138 138
FT CARBOHYD 154 154
FT CARBOHYD 158 158
FT CARBOHYD 197 197
FT CARBOHYD 234 234
FT CARBOHYD 241 241
FT CARBOHYD 262 262

```

FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 607 607 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 612 612 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 621 621 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 633 633 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 670 670 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 812 812 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 852 AA; 97203 MW; 2BB866345DEC915F CRC64;

Query Match 68.1%; Score 92; DB 1; Length 852;

Best Local Similarity 75.0%; Pred. No. 1.le-05; Mismatches 7; Indels 0; Gaps 0;

OY 1 RYIRVQACRAIRHIVRRIRROGLRRIL 28

DB 824 RAIEVQRAFRALHPRIRROGLRRIL 851

Search completed: August 14, 2002, 10:59:44
Job time: 500 sec

